



PCT09

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/936,883A

DATE: 01/26/2002

TIME: 12:40:03

Input Set : A:\F2-101DP1PCTsq.txt

Output Set: N:\CRF3\01262002\I936883A.raw

3 <110> APPLICANT: MIYATA, Toshio  
 5 <120> TITLE OF INVENTION: A Method for Detecting Megsin Protein and Use  
 6 Thereof  
 8 <130> FILE REFERENCE: F2-101DP1PCT  
 10 <140> CURRENT APPLICATION NUMBER: US/09/936,883A  
 11 <141> CURRENT FILING DATE: 2001-12-21  
 13 <150> PRIOR APPLICATION NUMBER: JP 1999-75305  
 14 <151> PRIOR FILING DATE: 1999-03-19  
 16 <150> PRIOR APPLICATION NUMBER: JP 1999-306623  
 17 <151> PRIOR FILING DATE: 1999-10-28  
 19 <160> NUMBER OF SEQ ID NOS: 21  
 21 <170> SOFTWARE: PatentIn Ver. 2.0  
 23 <210> SEQ ID NO: 1  
 24 <211> LENGTH: 1143  
 25 <212> TYPE: DNA  
 26 <213> ORGANISM: Homo sapiens  
 28 <220> FEATURE:  
 29 <221> NAME/KEY: CDS  
 30 <222> LOCATION: (1)..(1140)  
 32 <300> PUBLICATION INFORMATION:  
 33 <302> TITLE: A mesangium-predominant gene, megsin, is a new serpin  
 34 upregulated in IgA nephropathy.  
 35 <303> JOURNAL: J. Clin. Invest.  
 36 <304> VOLUME: 120  
 37 <305> ISSUE: 4  
 38 <306> PAGES: 828-836  
 39 <307> DATE: 1998-08-15  
 41 <400> SEQUENCE: 1  
 42 atg gcc tcc ctt gct gca gca aat gca gag ttt tgc ttc aac ctg ttc 48  
 43 Met Ala Ser Leu Ala Ala Asn Ala Glu Phe Cys Phe Asn Leu Phe  
 44 1 5 10 15  
 46 aga gag atg gat gac aat caa gga aat gga aat gtg ttc ttt tcc tct 96  
 47 Arg Glu Met Asp Asp Asn Gln Gly Asn Gly Asn Val Phe Phe Ser Ser  
 48 20 25 30  
 50 ctg agc ctc ttc gct gcc ctg gcc ctg gtc cgc ttg ggc gct caa gat 144  
 51 Leu Ser Leu Phe Ala Ala Leu Ala Leu Val Arg Leu Gly Ala Gln Asp  
 52 35 40 45  
 54 gac tcc ctc tct cag att gat aag ttg ctt cat gtt aac act gcc tca 192  
 55 Asp Ser Leu Ser Gln Ile Asp Lys Leu Leu His Val Asn Thr Ala Ser  
 56 50 55 60  
 58 gga tat gga aac tct tct aat agt cag tca ggg ctc cag tct caa ctg 240  
 59 Gly Tyr Gly Asn Ser Ser Asn Ser Gln Ser Gly Leu Gln Ser Gln Leu  
 60 65 70 75 80

ENTERED

## RAW SEQUENCE LISTING

DATE: 01/26/2002

PATENT APPLICATION: US/09/936,883A

TIME: 12:40:03

Input Set : A:\F2-101DP1PCTsq.txt

Output Set: N:\CRF3\01262002\I936883A.raw

```

62 aaa aga gtt ttt tct gat ata aat gca tcc cac aag gat tat gat ctc 288
63 Lys Arg Val Phe Ser Asp Ile Asn Ala Ser His Lys Asp Tyr Asp Leu
64      85      90      95
66 agc att gtg aat ggg ctt ttt gct gaa aaa gtg tat ggc ttt cat aag 336
67 Ser Ile Val Asn Gly Leu Phe Ala Glu Lys Val Tyr Gly Phe His Lys
68      100      105      110
70 gac tac att gag tgt gcc gaa aaa tta tac gat gcc aaa gtg gag cga 384
71 Asp Tyr Ile Glu Cys Ala Glu Lys Leu Tyr Asp Ala Lys Val Glu Arg
72      115      120      125
74 gtt gac ttt acg aat cat tta gaa gac act aga cgt aat att aat aag 432
75 Val Asp Phe Thr Asn His Leu Glu Asp Thr Arg Arg Asn Ile Asn Lys
76      130      135      140
78 tgg gtt gaa aat gaa aca cat ggc aaa atc aag aac gtg att ggt gaa 480
79 Trp Val Glu Asn Glu Thr His Gly Lys Ile Lys Asn Val Ile Gly Glu
80 145      150      155      160
82 ggt ggc ata agc tca tct gct gta atg gtg ctg gtg aat gct gtg tac 528
83 Gly Gly Ile Ser Ser Ser Ala Val Met Val Leu Val Asn Ala Val Tyr
84      165      170      175
86 ttc aaa ggc aag tgg caa tca gcc ttc acc aag agc gaa acc ata aat 576
87 Phe Lys Gly Lys Trp Gln Ser Ala Phe Thr Lys Ser Glu Thr Ile Asn
88      180      185      190
90 tgc cat ttc aaa tct ccc aag tgc tct ggg aag gca gtc gcc atg atg 624
91 Cys His Phe Lys Ser Pro Lys Cys Ser Gly Lys Ala Val Ala Met Met
92      195      200      205
94 cat cag gaa cgg aag ttc aat ttg tct gtt att gag gac cca tca atg 672
95 His Gln Glu Arg Lys Phe Asn Leu Ser Val Ile Glu Asp Pro Ser Met
96      210      215      220
98 aag att ctt gag ctc aga tac aat ggt ggc ata aac atg tac gtt ctg 720
99 Lys Ile Leu Glu Leu Arg Tyr Asn Gly Gly Ile Asn Met Tyr Val Leu
100 225      230      235      240
102 ctg cct gag aat gac ctc tct gaa att gaa aac aaa ctg acc ttt cag 768
103 Leu Pro Glu Asn Asp Leu Ser Glu Ile Glu Asn Lys Leu Thr Phe Gln
104      245      250      255
106 aat cta atg gaa tgg acc aat cca agg cga atg acc tct aag tat gtt 816
107 Asn Leu Met Glu Trp Thr Asn Pro Arg Arg Met Thr Ser Lys Tyr Val
108      260      265      270
110 gag gta ttt ttt cct cag ttc aag ata gag aag aat tat gaa atg aaa 864
111 Glu Val Phe Phe Pro Gln Phe Lys Ile Glu Lys Asn Tyr Glu Met Lys
112      275      280      285
114 caa tat ttg aga gcc cta ggg ctg aaa gat atc ttt gat gaa tcc aaa 912
115 Gln Tyr Leu Arg Ala Leu Gly Leu Lys Asp Ile Phe Asp Glu Ser Lys
116      290      295      300
118 gca gat ctc tct ggg att gct tct ggg ggt cgt ctg tat ata tca agg 960
119 Ala Asp Leu Ser Gly Ile Ala Ser Gly Gly Arg Leu Tyr Ile Ser Arg
120 305      310      315      320
122 atg atg cac aaa tct tac ata gag gtc act gag gag ggc acc gag gct 1008
123 Met Met His Lys Ser Tyr Ile Glu Val Thr Glu Glu Gly Thr Glu Ala
124      325      330      335
126 act gct gcc aca gga agt aat att gta gaa aag caa ctc cct cag tcc 1056

```

## RAW SEQUENCE LISTING

DATE: 01/26/2002

PATENT APPLICATION: US/09/936,883A

TIME: 12:40:03

Input Set : A:\F2-101DP1PCTsq.txt

Output Set: N:\CRF3\01262002\I936883A.raw

```

127 Thr Ala Ala Thr Gly Ser Asn Ile Val Glu Lys Gln Leu Pro Gln Ser
128      340      345      350
130 acg ctg ttt aga gct gac cac cca ttc cta ttt gtt atc agg aag gat 1104
131 Thr Leu Phe Arg Ala Asp His Pro Phe Leu Phe Val Ile Arg Lys Asp
132      355      360      365
134 gac atc atc tta ttc agt ggc aaa gtt tct tgc cct tga 1143
135 Asp Ile Ile Leu Phe Ser Gly Lys Val Ser Cys Pro
136      370      375      380
139 <210> SEQ ID NO: 2
140 <211> LENGTH: 380
141 <212> TYPE: PRT
142 <213> ORGANISM: Homo sapiens
144 <400> SEQUENCE: 2
145 Met Ala Ser Leu Ala Ala Ala Asn Ala Glu Phe Cys Phe Asn Leu Phe
146 1 5 10 15
148 Arg Glu Met Asp Asp Asn Gln Gly Asn Gly Asn Val Phe Phe Ser Ser
149 20 25 30
151 Leu Ser Leu Phe Ala Ala Leu Ala Leu Val Arg Leu Gly Ala Gln Asp
152 35 40 45
154 Asp Ser Leu Ser Gln Ile Asp Lys Leu Leu His Val Asn Thr Ala Ser
155 50 55 60
157 Gly Tyr Gly Asn Ser Ser Asn Ser Gln Ser Gly Leu Gln Ser Gln Leu
158 65 70 75 80
160 Lys Arg Val Phe Ser Asp Ile Asn Ala Ser His Lys Asp Tyr Asp Leu
161 85 90 95
163 Ser Ile Val Asn Gly Leu Phe Ala Glu Lys Val Tyr Gly Phe His Lys
164 100 105 110
166 Asp Tyr Ile Glu Cys Ala Glu Lys Leu Tyr Asp Ala Lys Val Glu Arg
167 115 120 125
169 Val Asp Phe Thr Asn His Leu Glu Asp Thr Arg Arg Asn Ile Asn Lys
170 130 135 140
172 Trp Val Glu Asn Glu Thr His Gly Lys Ile Lys Asn Val Ile Gly Glu
173 145 150 155 160
175 Gly Gly Ile Ser Ser Ala Val Met Val Leu Val Asn Ala Val Tyr
176 165 170 175
178 Phe Lys Gly Lys Trp Gln Ser Ala Phe Thr Lys Ser Glu Thr Ile Asn
179 180 185 190
181 Cys His Phe Lys Ser Pro Lys Cys Ser Gly Lys Ala Val Ala Met Met
182 195 200 205
184 His Gln Glu Arg Lys Phe Asn Leu Ser Val Ile Glu Asp Pro Ser Met
185 210 215 220
187 Lys Ile Leu Glu Leu Arg Tyr Asn Gly Gly Ile Asn Met Tyr Val Leu
188 225 230 235 240
190 Leu Pro Glu Asn Asp Leu Ser Glu Ile Glu Asn Lys Leu Thr Phe Gln
191 245 250 255
193 Asn Leu Met Glu Trp Thr Asn Pro Arg Arg Met Thr Ser Lys Tyr Val
194 260 265 270
196 Glu Val Phe Phe Pro Gln Phe Lys Ile Glu Lys Asn Tyr Glu Met Lys
197 275 280 285

```

## RAW SEQUENCE LISTING

DATE: 01/26/2002

PATENT APPLICATION: US/09/936,883A

TIME: 12:40:03

Input Set : A:\F2-101DP1PCTsq.txt

Output Set: N:\CRF3\01262002\I936883A.raw

```

199 Gln Tyr Leu Arg Ala Leu Gly Leu Lys Asp Ile Phe Asp Glu Ser Lys
200      290      295      300
202 Ala Asp Leu Ser Gly Ile Ala Ser Gly Gly Arg Leu Tyr Ile Ser Arg
203 305      310      315      320
205 Met Met His Lys Ser Tyr Ile Glu Val Thr Glu Glu Gly Thr Glu Ala
206      325      330      335
208 Thr Ala Ala Thr Gly Ser Asn Ile Val Glu Lys Gln Leu Pro Gln Ser
209      340      345      350
211 Thr Leu Phe Arg Ala Asp His Pro Phe Leu Phe Val Ile Arg Lys Asp
212      355      360      365
214 Asp Ile Ile Leu Phe Ser Gly Lys Val Ser Cys Pro
215      370      375      380

```

218 &lt;210&gt; SEQ ID NO: 3

219 &lt;211&gt; LENGTH: 29

220 &lt;212&gt; TYPE: DNA

221 &lt;213&gt; ORGANISM: Artificial Sequence

223 &lt;220&gt; FEATURE:

224 <223> OTHER INFORMATION: Description of Artificial Sequence:Artificially  
225 synthesized degenerative primer sequence

227 &lt;220&gt; FEATURE:

228 &lt;221&gt; NAME/KEY: misc\_feature

229 &lt;222&gt; LOCATION: 26, 29

230 &lt;223&gt; OTHER INFORMATION: n is a or g or c or t.

232 &lt;400&gt; SEQUENCE: 3

233 gtgaatgctg tgtacttaaa ggcaantgn

29

236 &lt;210&gt; SEQ ID NO: 4

237 &lt;211&gt; LENGTH: 17

238 &lt;212&gt; TYPE: DNA

239 &lt;213&gt; ORGANISM: Artificial Sequence

241 &lt;220&gt; FEATURE:

242 <223> OTHER INFORMATION: Description of Artificial Sequence:Artificially  
243 synthesized degenerative primer sequence

245 &lt;220&gt; FEATURE:

246 &lt;221&gt; NAME/KEY: misc\_feature

247 &lt;222&gt; LOCATION: 3, 9, 15

248 &lt;223&gt; OTHER INFORMATION: n is a or g or c or t.

250 &lt;400&gt; SEQUENCE: 4

251 aaagaaag gtcngc

17

254 &lt;210&gt; SEQ ID NO: 5

255 &lt;211&gt; LENGTH: 26

256 &lt;212&gt; TYPE: DNA

257 &lt;213&gt; ORGANISM: Artificial Sequence

259 &lt;220&gt; FEATURE:

260 <223> OTHER INFORMATION: Description of Artificial Sequence:Artificially  
261 synthesized degenerative primer sequence

263 &lt;220&gt; FEATURE:

264 &lt;221&gt; NAME/KEY: misc\_feature

265 &lt;222&gt; LOCATION: 6, 9, 12, 15, 18, 21

266 &lt;223&gt; OTHER INFORMATION: n is a or g or c or t.

## RAW SEQUENCE LISTING

DATE: 01/26/2002

PATENT APPLICATION: US/09/936,883A

TIME: 12:40:03

Input Set : A:\F2-101DP1PCTsq.txt

Output Set: N:\CRF3\01262002\I936883A.raw

268 <400> SEQUENCE: 5  
 269 atggcctcng cngcngcngc naaygc 26  
 272 <210> SEQ ID NO: 6  
 273 <211> LENGTH: 37  
 274 <212> TYPE: DNA  
 275 <213> ORGANISM: Artificial Sequence  
 277 <220> FEATURE:  
 278 <223> OTHER INFORMATION: Description of Artificial Sequence:Artificially  
 279 synthesized degenerative primer sequence  
 281 <400> SEQUENCE: 6  
 282 cgacctccag aggcaattcc agagagatca gccctgg 37  
 285 <210> SEQ ID NO: 7  
 286 <211> LENGTH: 34  
 287 <212> TYPE: DNA  
 288 <213> ORGANISM: Artificial Sequence  
 290 <220> FEATURE:  
 291 <223> OTHER INFORMATION: Description of Artificial Sequence:Artificially  
 292 synthesized degenerative primer sequence  
 294 <400> SEQUENCE: 7  
 295 gtcttccaag cctacagatt tcaagtggct cctc 34  
 298 <210> SEQ ID NO: 8  
 299 <211> LENGTH: 30  
 300 <212> TYPE: DNA  
 301 <213> ORGANISM: Artificial Sequence  
 303 <220> FEATURE:  
 304 <223> OTHER INFORMATION: Description of Artificial Sequence:Artificially  
 305 synthesized antisense primer sequence  
 307 <400> SEQUENCE: 8  
 308 gctcagggca gtgaagatgc tcagggaaga 30  
 311 <210> SEQ ID NO: 9  
 312 <211> LENGTH: 27  
 313 <212> TYPE: DNA  
 314 <213> ORGANISM: Artificial Sequence  
 316 <220> FEATURE:  
 317 <223> OTHER INFORMATION: Description of Artificial Sequence:Artificially  
 318 synthesized antisense primer sequence  
 320 <400> SEQUENCE: 9  
 321 ctgacgtgca cagtcacctc gagcacc 27  
 324 <210> SEQ ID NO: 10  
 325 <211> LENGTH: 36  
 326 <212> TYPE: DNA  
 327 <213> ORGANISM: Artificial Sequence  
 329 <220> FEATURE:  
 330 <223> OTHER INFORMATION: Description of Artificial Sequence:Artificially  
 331 synthesized sense primer sequence  
 333 <400> SEQUENCE: 10  
 334 gaggtctcag aagaaggcac tgaggcaact gctgcc 36  
 337 <210> SEQ ID NO: 11  
 338 <211> LENGTH: 15

Use of n and/or Xaa has been detected in the Sequence Listing.  
 Review the Sequence Listing to insure a corresponding  
 explanation is presented in the <220> to <223> fields of  
 each sequence using n or Xaa.

## VERIFICATION SUMMARY

DATE: 01/26/2002

PATENT APPLICATION: US/09/936,883A

TIME: 12:40:04

Input Set : A:\F2-101DPI\CTsq.txt

Output Set: N:\CRF3\01262002\I936883A.raw

L:10 M:270 C: Current Application Number differs, Replaced Application Number

L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:233 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3

L:251 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4

L:269 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5

L:464 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:18

L:465 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:18

L:472 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:18

L:473 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:18

L:573 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:19

L:579 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:19